Query Expansion with Semantic-based Ellipsis Reduction for Conversational IR

ASCFDA
Chia-Yuan Chang, Hsien-Hao Chen, Ning Chen,
Wei-Ting Chiang, Chih-Hen Lee, Yu-Hsuan Tseng

Speaker: Hsien Hao, Chen

Agenda

1. Introduction
2. Coreference Query Reformation (CQR)
3. Semantic-based Ellipsis Reduction (SER)
4. Retrieve & Rerank
5. Manually rewritten utterance
6. Results & Conclusion

1. Introduction

“You're a wizard, Harry.”

1-1. Problem Definition

- Topic_number: 83

Raw Query	
83-1	What are some interesting facts about **bees**
83-2	Why doesn't $\boxed{\text{it}}$ spoil ?
83-3	Why are so many dying ? subject missing
83-4	What can be done to stop $\boxed{\text{it}}$?
83-5	What has happened to $\boxed{\text{their}}$ habitat?

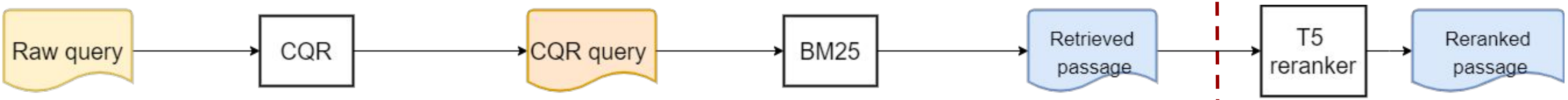
1-2. Query Ambiguity

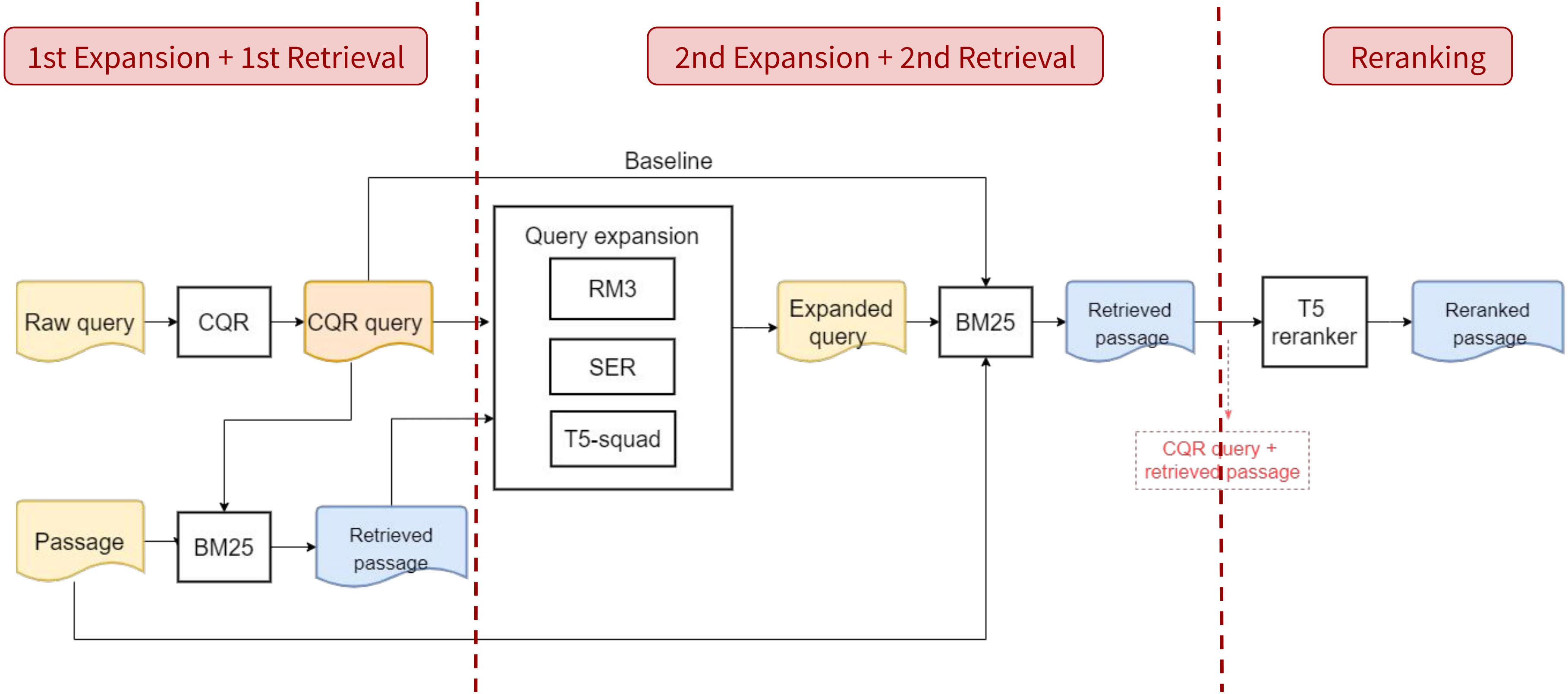
- In a conversational system, the semantic ambiguity may come from:
  1. Expressing the same thing by various kinds of words
  2. Pronoun usage
  3. Omitting the repeating subjects
- The potential sources of supplement information to fix it:
  1. Historical queries
  2. Highly correlated passages

1-3. Pipeline (Baseline)

1st expansion + 1st retrieve

Rerank



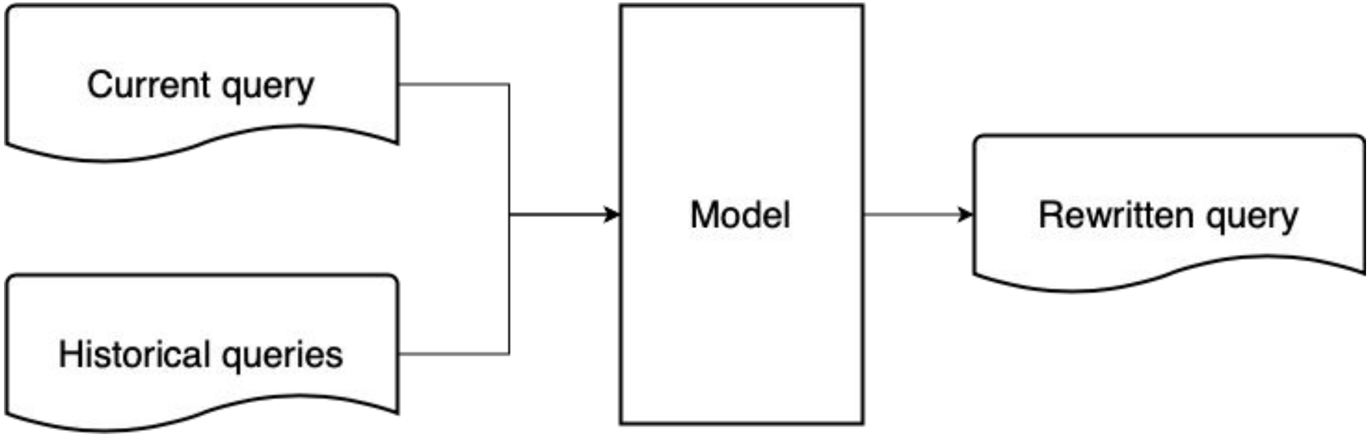


1st Expansion + 1st Retrieval
2nd Expansion + 2nd Retrieval
Reranking
Baseline
Query expansion
RM3
SER
T5-squad
Raw query
CQR
CQR query
Expanded query
BM25
Retrieved passage
T5 reranker
Reranked passage
CQR query + retrieved passage
Passage
BM25
Retrieved passage

2. Coreference Query Reformation (CQR)

“Oh yes, the past can hurt. But you can either run from it, or learn from it.”

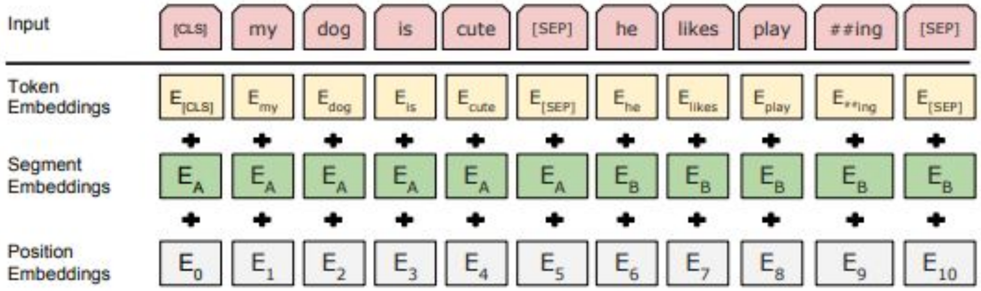
2-1. Coreference Query Reformation



2-2. Introduction of Transformers

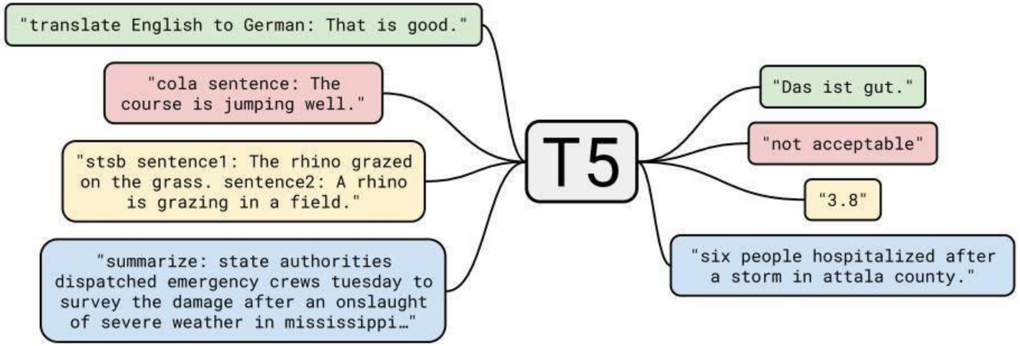
BERT

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.



T5

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. arXiv preprint arXiv:1910.10683.



2-3. CQR model (T5-CQR)

- Pretrained model: T5-based
- Fine-tuning
  - Dataset: CANARD
  - Dialog with rewritten questions
- Inference
  - Input:
    - only use a query of each turn
  - Output:
    - rewritten query of the last one in input

Input	
Metadata	"Ara Parseghian", "First national title", "When did Ara Parseghian's win his first title.", "In 1966,"
Question query$_2$	"What was their record for that year?",

Output	
Rewrite **query$_2$***	"What was Ara Parseghian's record for 1966?",

Let's see how fantastic the CQR is...

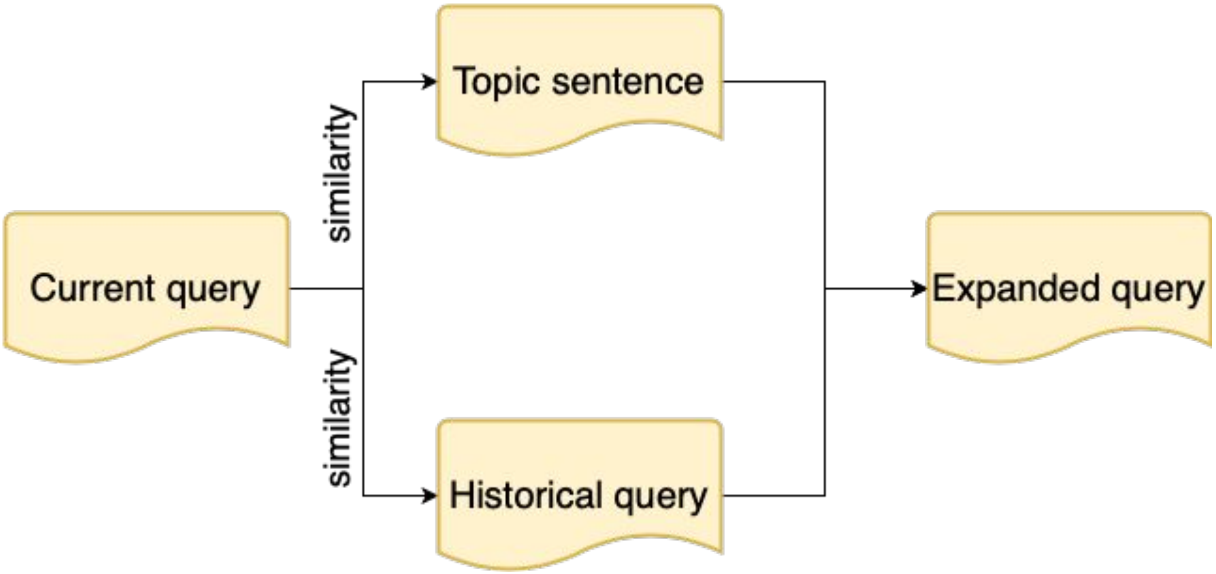
Raw Query	
83-1	What are some interesting facts about **bees**
83-2	Why doesn't **it** spoil ?
83-3	Why are so many dying ?
83-4	What can be done to stop **it** ?
83-5	What has happened to **their** habitat?

CQR Query	
83-1	What are some interesting facts about **bees**
83-2	Why doesn't **bees** spoil ?
83-3	Why are so many **dying from bees** ?
83-4	What can be done to stop **bees dying** ?
83-5	What has happened to **bees** habitat ?

3. Semantic-based Ellipsis Reduction (SER)

“Whoever you are— I have always depended on the kindness of strangers.”

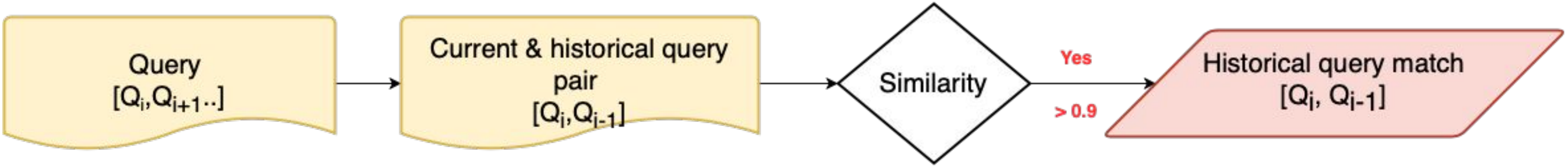
3-2. SER Model


Topic sentence
similarity
Current query
Expanded query
similarity
Historical query

3-3. Historical Queries

- Transformer model: roberta-large
  - Trained on NLI + STSb
- Cosine Similarity Threshold: 0.8 ~ 0.9
  - Too high similarity is meaningless

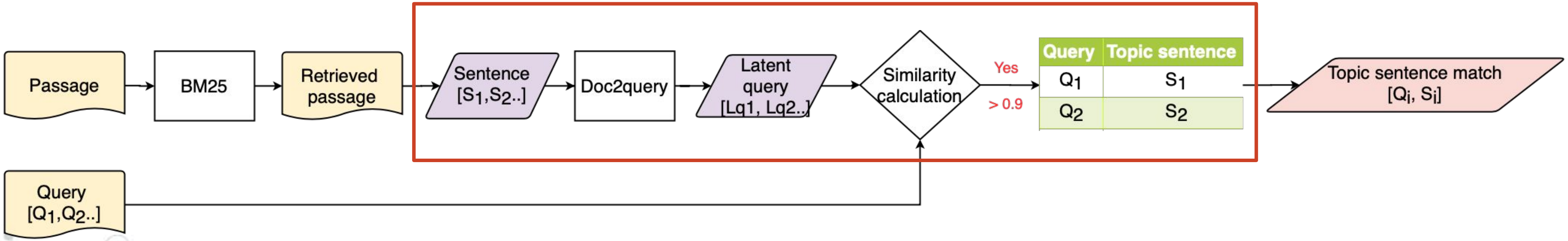
historical	83-3	Why are so many **dying from bees** ?
current	83-4	What can be done to stop **bees dying** ?



ASCFDA

3-4. Topic Sentence

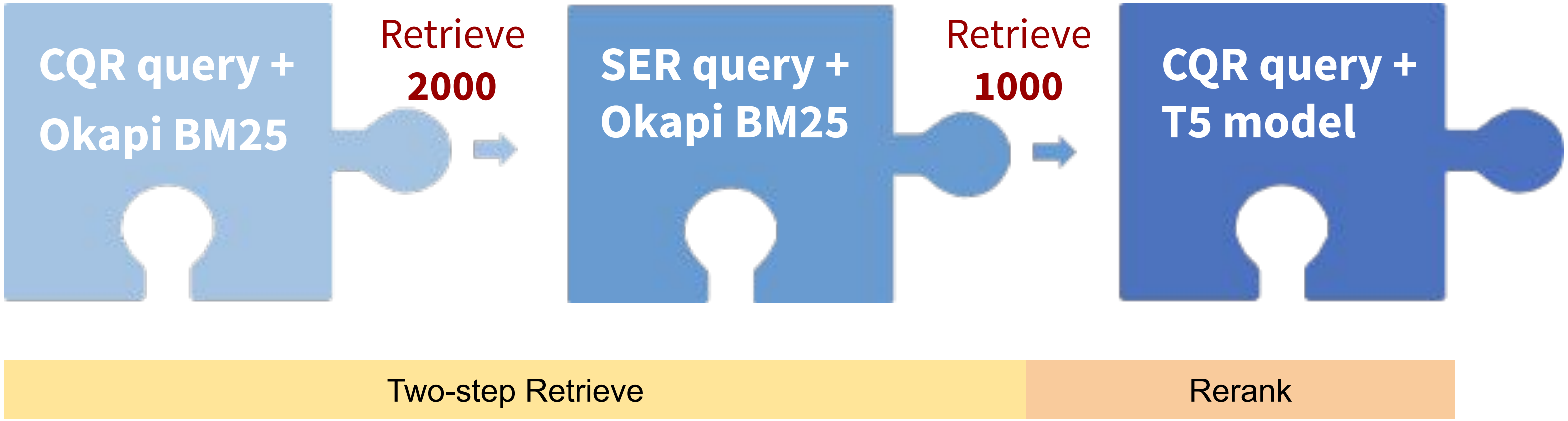
- Sentence + doc2query → latent query
- Transformer model: roberta-large
  - trained on NLI+STSb
- Cosine Similarity Threshold: 0.9



4. Retrieve & Rerank

“You will ride eternal, shiny and chrome.”

4-1. Retrieve & Rerank



4-1. Retrieve & Rerank

- Two-step Retrieve:
  - CQR query + Okapi BM25 -> retrieve 2000.
  - SER query + Okapi BM25 -> retrieve 1000.

- Rerank:
  - CQR query + T5 model

5. Manually rewritten utterance

“Just keep swimming.”

5-1. T5-SQuAD for Query Expansion

- Purpose
  - To extract the keywords of manual responses with queries information
- Method description
  - Use the **manual result** as the content of T5-squad pretrained model
  - And ask the **CQR utterance** to T5-SQuAD pretrained model
  - Expand CQR utterances with the answer of T5-squad pretrained model

SQuAD Example

Paragraph:
...These later laws had a low cost to society–the species were relatively rare–and little opposition was raised.
Question:
Which laws faced significant opposition?
Answer:
Later laws

T5-SQuAD for QE

Paragraph:
Manual_response_1
Question:
Query_1
Answer:
The QE materials

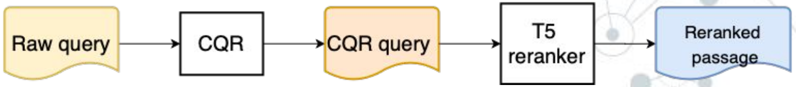
QE: Query_1 + QE materials

6. Results & Conclusion

Do Androids Dream of Electric Sheep?

6-1. Results

- Baseline model performs the best.
- QE(RM3) and QE(SER) reach higher recalls than baseline.



	Trec 2019					Trec 2020				
Stage	Retrieve			Rerank		Retrieve + Rerank				
Raw utterance only	mAP@1000	R@1000	R@2000	mAP@1000	R@1000	mAP@1000	R@1000	NDCG@3	NDCG@5	NDCG@1000
Raw queries	0.1077	0.4182	0.4681	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Baseline	0.2497	0.7628	0.8260	**0.3724**	0.8060	**0.3096**	0.6106	**0.4579**	**0.4472**	0.4943
QE(RM3)	**0.2845**	**0.8024**	**0.8563**	0.3695	**0.8252**	0.3092	**0.6405**	0.4511	0.4362	**0.5003**
QE(SER)	0.2434	0.7674	0.8288	0.3713	0.8089	0.3090	0.6131	0.4576	0.4456	0.4934
Manually rewritten utterance	mAP@1000	R@1000	R@2000	mAP@1000	R@1000	mAP@1000	R@1000	NDCG@3	NDCG@5	NDCG@1000
QE(T5-squad)	N/A	N/A	N/A	N/A	N/A	0.3102	0.6498	0.4663	0.4514	0.5131

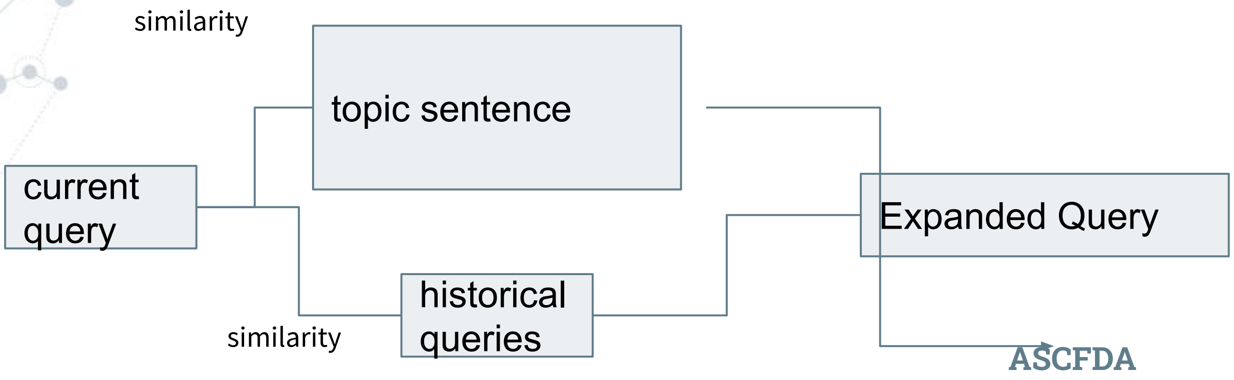
6-2. Conclusion

- The mismatch between queries and documents is crucial in a conversational task.
- The considerable potential of a semantic-based relevance-feedback method.
- T5 domination. “When in doubt, C4!”

Cheers. Thank you for your Attention!

Result of the SER

CQR Query	Top Sentence Extracted by SER
95-6 Tell me more about biodegradable plastics.	Biodegradable plastics are plastics that decompose by the action of living organisms, usually bacteria
102-9 How much of an increase is there in social security?	How much faster will it grow as a share of the economy? Social Security benefits amounted to 4.9 percent of GDP in 2014
88-4 Why was slavery important? (in the Ottman Empire)	Why was slavery so important to the American South during the period near the Civil War? ...


similarity
topic sentence
current query
Expanded Query
historical queries
similarity
ASCFDA